IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Group Art Unit: *Unassigned*
Examiner: *Unassigned*

In Re PATENT APPLICATION of:

Applicant(s)	:	Sayori SHIMOHATA, et al	
Application No.	:	*To Be Assigned*	
Filed	:	*Concurrently*	Customer No. 26694
For	:	DOCUMENT PROCESSING DEVICE AND DOCUMENT PROCESSING METHOD	Patent Trademark Office
Attorney Docket	:	31759-190464	June 23, 2003

SUMBISSION OF PRIORITY DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Submitted herewith is a certified copy of Application No. JP2002-182883 filed on June 24, 2002 in Japan, the priority of which is claimed in the present application under the provisions of 35 U.S.C. 119.

Respectfully submitted,

Norman N. Kunitz
Registration No. 20,586
VENABLE LLP
P.O. Box 34385
Washington, D.C. 20043-9998
Telephone: (202) 962-4800
Telefax: (202) 962-8300

Reg. No. 33,074

NNK/elw
#465624

日　本　国　特　許　庁
JAPAN　PATENT　OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日
Date of Application:　２００２年　６月２４日

出願番号
Application Number:　特願２００２－１８２８８３

[ST.10/C]:　[ＪＰ２００２－１８２８８３]

出願人
Applicant(s):　沖電気工業株式会社

Inventor: Sayori SHIMOKAWA, et al
ATTY DKT: 31759-190464
Customer NO: 26694

２００３年　４月　１日

特許庁長官
Commissioner,
Japan Patent Office　太田信一郎

出証番号　出証特２００３－３０２２４３９

【書類名】　特許願

【整理番号】　KN002523

【提出日】　平成14年 6月24日

【あて先】　特許庁長官　及川　耕造　殿

【国際特許分類】　G06F 15/00

【発明者】

【住所又は居所】　東京都港区虎ノ門１丁目７番１２号　沖電気工業株式会社内

【氏名】　下畑　さより

【発明者】

【住所又は居所】　東京都港区虎ノ門１丁目７番１２号　沖電気工業株式会社内

【氏名】　池野　篤司

【特許出願人】

【識別番号】　000000295

【氏名又は名称】　沖電気工業株式会社

【代表者】　篠塚　勝正

【代理人】

【識別番号】　100090620

【弁理士】

【氏名又は名称】　工藤　宣幸

【手数料の表示】

【予納台帳番号】　013664

【納付金額】　21,000円

【提出物件の目録】

【物件名】　明細書　1

【物件名】　図面　1

【物件名】　要約書　1

【包括委任状番号】　9006358

【プルーフの要否】　　要

【書類名】　明細書

【発明の名称】　文書処理装置および方法

【特許請求の範囲】

【請求項１】　文字情報を含む複数の文書を要素とする集合に関して処理を行う文書処理装置において、

前記集合中の各文書に関し、前記文字情報の共通性を抽出して、前記集合全体に共通の意味内容を表現した文書である共通文書を生成する共通文書生成手段を備えることを特徴とする文書処理装置。

【請求項２】　請求項１の文書処理装置において、

前記共通文書生成手段は、

前記集合中の複数の文書をもとに所定の生成手順を実行して、新たな文書として前記共通文書を生成するか、または、

予め文字情報に共通性のある文書を選んで前記集合を構成した上で、前記集合中の複数の文書のなかから所定の選択手順に応じて１つの文書を選択し、選択した当該文書を前記共通文書とすることで前記共通文書を生成することを特徴とする文書処理装置。

【請求項３】　請求項２の文書処理装置において、

前記選択手順では、

前記集合中の各文書に共通して、所定のしきい値以上、出現する頻出表現を検出し、当該頻出表現が最も多く含まれる文書を選択することを特徴とする文書処理装置。

【請求項４】　請求項２の文書処理装置において、

前記共通文書と、前記集合中の各文書との差分となる文字情報である差分文字情報を抽出する差分文字情報抽出手段、または、

前記共通文書と、前記集合中の各文書との共通の文字情報である共通文字情報を抽出する共通文字情報抽出手段を備えることを特徴とする文書処理装置。

【請求項５】　請求項４の文書処理装置において、

前記文書を画面表示する表示手段と、

前記集合の要素である各文書に論理構造を付与し、当該論理構造の付与に際して、少なくとも前記差分文字情報または共通文字情報のいずれかに関しては、その旨を明示する論理構造付与手段と、

当該論理構造付与手段で論理構造を付与した後の各文書を蓄積する文書蓄積手段とを備え、

前記表示手段による画面表示では、当該論理構造を利用した画面表示を行うことを特徴とする文書処理装置。

【請求項６】　請求項５の文書処理装置において、

前記表示手段は、

前記共通文書と当該共通文書に対する各文書の差分文字情報とを、ユーザからの操作に応じて画面表示することを特徴とする文書処理装置。

【請求項７】　請求項６の文書処理装置において、

前記表示手段は、

少なくとも前記差分文字情報を含む文書に関し、オンラインまたはオフラインの出所を示す出典情報を、ユーザからの操作に応じて画面表示することを特徴とする文書処理装置。

【請求項８】　請求項７の文書処理装置において、

前記表示手段は、

ユーザが画面表示された出典情報に対して所定の操作を行うと、前記論理構造を利用して、その出典情報に対応する文書と、前記共通文書との差分文字情報または共通文字情報を識別する所定の識別表示を、画面表示される共通文書に対して実行することを特徴とする文書処理装置。

【請求項９】　文字情報を含む複数の文書を要素とする集合に関して処理を行う文書処理方法において、

共通文書生成手段が、前記集合中の各文書に関し、前記文字情報の共通性を抽出して、前記集合全体に共通の意味内容を表現した文書である共通文書を生成することを特徴とする文書処理方法。

【請求項１０】　請求項９の文書処理方法において、

前記共通文書生成手段は、

前記集合中の複数の文書をもとに所定の生成手順を実行して、新たな文書として前記共通文書を生成するか、または、

予め文字情報に共通性のある文書を選んで前記集合を構成した上で、前記集合中の複数の文書のなかから所定の選択手順に応じて１つの文書を選択し、選択した当該文書を前記共通文書とすることで前記共通文書を生成することを特徴とする文書処理方法。

【請求項１１】　請求項１０の文書処理方法において、

前記選択手順では、

前記集合中の各文書に共通して、所定のしきい値以上、出現する頻出表現を検出し、当該頻出表現が最も多く含まれる文書を選択することを特徴とする文書処理方法。

【請求項１２】　請求項１０の文書処理方法において、

前記共通文書と、前記集合中の各文書との差分となる文字情報である差分文字情報を抽出するか、または、

前記共通文書と、前記集合中の各文書との共通の文字情報である共通文字情報を抽出することを特徴とする文書処理方法。

【請求項１３】　請求項１２の文書処理方法において、

論理構造付与手段が、前記集合の要素である各文書に論理構造を付与し、当該論理構造の付与に際して、少なくとも前記差分文字情報または共通文字情報のいずれかに関しては、その旨を明示し、

文書蓄積手段が、当該論理構造付与手段で論理構造を付与した後の各文書を蓄積し、

表示手段は、当該論理構造を利用した画面表示を行うことを特徴とする文書処理方法。

【請求項１４】　請求項１３の文書処理方法において、

前記表示手段は、

前記共通文書と当該共通文書に対する各文書の差分文字情報とを、ユーザからの操作に応じて画面表示することを特徴とする文書処理方法。

【請求項１５】　請求項１４の文書処理方法において、

前記表示手段は、

少なくとも前記差分文字情報を含む文書に関し、オンラインまたはオフラインの出所を示す出典情報を、ユーザからの操作に応じて画面表示することを特徴とする文書処理方法。

【請求項１６】　請求項１５の文書処理方法において、

前記表示手段は、

ユーザが画面表示された出典情報に対して所定の操作を行うと、前記論理構造を利用して、その出典情報に対応する文書と、前記共通文書との差分文字情報または共通文字情報を識別する所定の識別表示を、画面表示される共通文書に対して実行することを特徴とする文書処理方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は文書処理装置および文書処理方法に関し、例えば、同じキーワードを検索キーとしてテキストデータベースを検索した結果として得られる複数の同類のテキストを処理し、表示する場合などに適用して好適なものである。

【０００２】

【従来の技術】

従来のこの種の装置としては次の文献１に開示されるものがある。

文献１：特開平９－２３１２３８号公報

文献１の表示装置が実行する処理は、テキスト集合を自動的に複数個のグループに分割する分割ステップと、当該分割ステップによって得られた各グループのそれぞれについて、当該グループの属性を表現する主題分類情報を生成する生成ステップと、当該生成ステップで求めた各グループの主題分類情報をグループ別に区分して表示する表示ステップとから構成されている。

ここで、主題分類情報とは、テキストの内容に対応した情報で、キーワードの組あるいは、短い文章を指す。

文献１の表示装置ではまた、前記グループと検索条件の間の適合度、および、グループ内の各テキストの、当該グループに対する所属度を算出するステップを有

し、これらの値にしたがって表示するグループやテキストの順番を選択することも可能である。

【０００３】

【発明が解決しようとする課題】

しかしながら、上記のような表示装置では、グループごとに提示される各グループの主題分類情報、すなわち、キーワードの組や短い文章からそのグループに含まれるテキストの内容を判断しなければならない。多くの場合、キーワードの組や短い文章だけからそのグループに含まれるテキストの内容（あるいは、グループの概要）を的確に判断することは困難であるため、結局、ユーザはグループに含まれる個々のテキストを読むことによってしか、検索結果を確認することができず、グループの概要を知ることもできない。

【０００４】

したがって、検索結果を確認したり、グループの概要を知るために長い時間と手数を要し、利便性に欠ける構成となっている。

【０００５】

また、上記表示装置において当該主題分類情報が得られるのは、テキスト集合が得られ、さらに当該テキスト集合を分割してグループが得られたあとであるから、テキスト集合が得られた時点では当該主題分類情報さえ存在せず、ユーザがテキスト集合の概要を知るには、個々のテキストを読む以外の方法はなく、極めて利便性が低い。

【０００６】

【課題を解決するための手段】

かかる課題を解決するために、第１の発明では、文字情報を含む複数の文書を要素とする集合に関して処理を行う文書処理装置において、前記集合中の各文書に関し、前記文字情報の共通性を抽出して、前記集合全体に共通の意味内容を表現した文書である共通文書を生成する共通文書生成手段を備えることを特徴とする。

【０００７】

また、第２の発明では、文字情報を含む複数の文書を要素とする集合に関して

処理を行う文書処理方法において、共通文書生成手段が、前記集合中の各文書に関し、前記文字情報の共通性を抽出して、前記集合全体に共通の意味内容を表現した文書である共通文書を生成することを特徴とする。

【０００８】

【発明の実施の形態】

（Ａ）実施形態

以下、本発明にかかる文書処理装置および方法を、検索エンジンを含む閲覧システムに適用した場合を例に、実施形態について説明する。

【０００９】

（Ａ－１）第１の実施形態の構成

本実施形態にかかる閲覧システム１０の全体構成例を、図１に示す。図１の各構成要素１～５は、イントラネット内や、あるいは１つの情報処理装置の内部に配置されるものであってもよいが、ここでは、インターネット上に分散配置されるものとして説明する。

【００１０】

図１において、当該閲覧システム１０は、入出力部１と、テキストデータベース２と、検索エンジン３と、テキスト集合蓄積部４と、テキスト加工処理部５と、作業用データベース６とを備えている。

【００１１】

このうち入出力部１は、当該閲覧システム１０を利用するユーザＵ１の操作する通信端末に相当する部分で、ハードウエア的には例えばキーボードやマウスなどのポインティングデバイス、ディスプレイ装置、ハードディスクやメモリ装置などを有するパーソナルコンピュータ等が該当し、ソフトウエア的には、当該パーソナルコンピュータが搭載するブラウザ等が該当し得る。

【００１２】

ブラウザとしてはＷｅｂページを閲覧するためのＷｅｂブラウザがよく知られているが、単にブラウザと呼ぶときは、必ずしもＷｅｂブラウザにかぎらず、何らかの情報を閲覧する機能を持つソフトウエア全般を指す。

【００１３】

検索エンジン３は、ユーザＵ１の操作に応じて入出力部１から供給される１または複数のキーワードをもとに全文検索を実行する部分である。

【００１４】

全文検索とは文書中のすべての文字列から目的の文字列を検索する操作をいう。したがって、例えば、新聞の内容を記述したＷｅｂページを検索する場合には、当該Ｗｅｂページを構成するＨＴＭＬファイル中の全文字列が検索の対象になる。

【００１５】

全文検索の機能は、必要ならば前記入出力部１を有するパーソナルコンピュータに搭載してもかまわないが、Ｗｅｂ（WWW）上ならば、専門の検索サービス業者がすでに提供している検索サービスを利用することができる。

【００１６】

テキストデータベース２は、ハードウエア的には、ハードディスクや光ディスクなどの記憶装置を利用して、多数のテキストを蓄積しているデータベースである。ここで、テキストとは、文書（ドキュメント）と同義である。文書には、データ形式がテキスト形式であるテキストデータ（プレーンテキスト形式のデータ）のほか、ＧＩＦやＪＰＥＧ等の画像データなども含まれ得る。通常の１つのＷｅｂページは、基本となる１つのＨＴＭＬファイル（データ形式として、ＨＴＭＬ形式はテキスト形式の一種である）のほかに、１または複数の画像ファイルなどによって構成され得るので、この文書に該当し得る。

【００１７】

この意味で、テキストデータベース２は、各種のＷｅｂページを提供する１または複数のＷｅｂサーバと見ることができる。

【００１８】

また、Ｗｅｂ上の検索サービス業者が検索の対象としているのは、世界中のＷｅｂページであることからすると、テキストデータベース２は、世界中に分散配置された膨大な数のＷｅｂページ（Ｗｅｂサーバ）によって構成されるＷｅｂ（ワールド・ワイド・ウエブ）そのものであると見ることもできる。

【００１９】

もちろん、テキストデータベース２は、テキスト（文書）を蓄積するデータベースであるから、Ｗｅｂページ以外の文書（例えば、ＸＭＬで記述された文書や、ＰＤＦなどの電子出版用のデータ形式で記述された文書など）が含まれていてもかまわない。

【００２０】

ＨＴＭＬ形式では、文字の位置や大きさなどを情報の送り手側が詳細に指定することが難しく、色彩の表現力などの点でも、通常の紙媒体の出版物（雑誌や書籍など）に比べるとかなり劣るため、インターネット上の出版物には、送り手側の意思をより忠実に反映することが可能なＰＤＦ形式などが利用されることが多い。なお、ＰＤＦ形式で記述された文書は、通常のＷｅｂブラウザの機能だけでは閲覧できないため、入出力部１が通常のＷｅｂブラウザだけしか搭載していない場合には、Ｗｅｂブラウザの機能を拡張するプラグインソフトを搭載することが必要になる。

【００２１】

ＰＤＦ形式など、通常のテキスト形式と異なるデータ形式で記述されたファイルは、検索の対象とする前にテキスト形式に変換しておくこと等により、容易に、検索エンジン３の検索対象とすることができる。

【００２２】

また、画像データとして文字が記述されることもあり得るが、このような文字も、適宜、テキスト形式に変換することによって検索エンジン３による検索の対象とすることが可能である。

【００２３】

テキスト加工処理部５は、検索エンジン３が前記キーワードを用いた検索の結果として得た複数の文書を加工する部分で、加工後の文書は、テキスト集合蓄積部４に蓄積する。本実施形態では、検索エンジン３による検索の結果として、内容の類似した複数の文書が得られる場合を想定する。具体的には、例えば、同一の事件に関して記述した同日付けの異なる新聞社による新聞記事などは、ここでいう内容の類似した複数の文書に該当し得る。

【００２４】

一般的には、１つの検索に関し、検索エンジン３に供給するキーワードの数が多いほど、また個々のキーワードが特徴的で識別性が高いものであるほど、検索結果として得られる複数の文書の内容は類似したものとなる傾向がある。検索の結果として得られる文書の数は偶発的で予測困難な事象であるから、１つの文書しか得られない可能性もあるが、テキストデータベース２に蓄積されている文書の数が十分に多ければ、多くの場合、複数の文書が得られる。

【００２５】

本実施形態では、検索エンジン３による検索の結果として得られた内容の類似した複数の文書は、１つのテキスト集合（文書集合）を構成するものと考え、当該テキスト集合をテキスト加工処理部５の処理の対象とする。なお、当該テキスト集合は、前記文献１の用語との関係では、前記グループではなく、前記テキスト集合に相当する概念である。

【００２６】

（Ａ－１－１）テキスト加工処理部の内部構成例

図１に示すように、当該テキスト加工処理部５は、主題情報生成部５Ａと、差分情報生成部５Ｂと、情報提示部５Ｃとを備えている。

【００２７】

このうち主題情報生成部５Ａは、１つのテキスト集合中の全文書の内容をもとに主題情報を生成する部分である。主題情報とは、当該テキスト集合の主題を示すのに十分な内容を備えた文章である。テキスト集合の主題は、基本的に、１つのテキスト集合中の全文書に共通する内容の文章によって表現される。

【００２８】

例えば、１つのテキスト集合ＴＸＧ１が３つの文書ＴＸ１～ＴＸ３から構成されている場合、テキスト集合ＴＸＧ１の主題情報ＴＨ１は、文書ＴＸ１～ＴＸ３のすべてに共通する内容の文章として表現することができる。

【００２９】

本実施形態における主題情報ＴＨ１の表現法には大きく分けて２通りの方法がある。その１つは、文書ＴＸ１～ＴＸ３の内容をもとに、これらの要約となる新たな文書ＴＸＡを生成し当該文書（要約）ＴＸＡによって主題情報ＴＨ１を表現

する方法（要約生成法）であり、もう１つは、文書ＴＸ１～ＴＸ３のなかから適切な文書を選択し、選択した文書（例えば、ＴＸ３）自体で主題情報ＴＨ１を表現する方法（代表選択法）である。

【００３０】

要約生成法の実現には、例えば、前記文書ＴＸ１～ＴＸ３に共通する文節を検出し、検出された各文節を組み合わせることによって前記要約ＴＸＡを生成する方法など、様々な方法が使用可能であるが、一例として、次の文献２に記載された方法を用いることもできる。

【００３１】

文献２：Columbia Multi-document Summarization: Approach and Evaluation

K.R. McKeown, R. Barzilay, D. Evans, V. Hatzivassiloglou, M.Y. Kan, B. Schiffman, S. Teufel　DUC'01

代表選択法の実現にも様々な方法が使用可能であるが、例えば、前記文書ＴＸ１～ＴＸ３に共通して数多く出現する表現（頻出表現）を抽出し、文書ＴＸ１～ＴＸ３のうち当該頻出表現が最も多く含まれる文書（例えば、ＴＸ３）を代表として選択する方法を用いることができる。

【００３２】

差分情報生成部５Ｂは、前記主題情報ＴＨ１に対する各文書（要約生成法を用いた場合にはＴＸ１～ＴＸ３、代表選択法を用いた場合には、代表として選択された以外の２つの文書（例えば、ＴＸ１，ＴＸ２））の差分を抽出する部分である。当該差分（差分情報）は、前記頻出表現を含む単位を各文書に共通する共通情報とし、頻出表現を含まない単位を各文書に固有な固有情報とすれば、当該固有情報として抽出される。ここで、単位とは、節、文、段落などの文法上の単位を指す。

【００３３】

差分を抽出したあと、各文書中の差分に該当する単位をマークアップ言語のタグの属性によって、当該単位が差分に該当する旨を指定することができる。

【００３４】

例えば、当該マークアップ言語がＸＭＬ（データ形式として、ＸＭＬ形式はテキスト形式の一種である）の場合には、開始タグと終了タグで当該単位を挟み、開始タグの中に記述する属性によって、当該単位が差分に該当する旨を記述することができる。この場合、必要に応じて、差分情報生成部５Ｂにおいて、ＨＴＭＬなどからＸＭＬへのデータ形式の変換を実行することになる。当該単位が差分に該当する旨を示し、なおかつ再利用可能な形で保存するためには、もともとテキストデータベース２上の文書がＸＭＬ文書でそのようなタグや属性がすでに定義されている場合などを除き、通常は、新たなタグや、新たな属性の定義が必要になり、このような定義が許容されるＸＭＬ形式を利用する必要があるからである。

【００３５】

前記文書ＴＸ１～ＴＸ３を当該ＸＭＬ形式に変換したあとの文書をＸＸ１～ＸＸ３と書く。ＸＭＬ形式の文書ＸＸ１は前記ＴＸ１に対応し、ＸＭＬ形式の文書ＸＸ２は前記ＴＸ２に対応し、ＸＭＬ形式の文書ＸＸ３は前記ＴＸ３に対応する。

【００３６】

ただしＸＭＬ文書は、タグを用いて文書の論理構造を示すだけなので、実際に、各ＸＭＬ文書ＸＸ１～ＸＸ３の入出力部１における表示方法（ユーザＵ１が各文書を閲覧する場合の見え方（すなわち、スタイル））を定義するには、スタイルシート言語を用いて具体的な表示方法を定義する必要がある。

【００３７】

情報提示部５Ｃは、前記主題情報生成部５Ａで得られた主題情報ＴＨ１、差分情報生成部５Ｂで得られたＸＭＬ文書ＸＸ１～ＸＸ３などを、入出力部１のブラウザで表示するのに適した所定の表示形態に加工してユーザＵ１に提示する部分である。

【００３８】

したがって、前記スタイルシート言語を用いた表示方法の定義も、この情報提示部５Ｃで行うようにするとよい。

【００３９】

具体的な表示方法については予め決定し、当該表示方法に対応するスタイルシート言語を、当該情報提示部５Ｃに付与しておけば、情報提示部５Ｃは、前記主題情報ＴＨ１やＸＭＬ文書ＸＸ１～ＸＸ３などが与えられたとき、自動的に、前記表示形態に加工することができる。

【００４０】

前記テキスト集合蓄積部４には、スタイルシート言語によって表示方法まで具体的に定義されたＸＭＬ文書ＸＸ１～ＸＸ３を蓄積しておくための記憶装置である。テキスト集合蓄積部４としては、前記入出力部１を有する通信端末が搭載したハードディスクなどの一部の記憶領域を利用してもよいが、インターネット上でオンラインストレージサービスを提供する事業者の持つストレージサーバなどを利用することもできる。

【００４１】

いずれにしても、主題情報生成部５Ａ、差分情報生成部５Ｂ、情報提示部５Ｃなどで行う各処理は、著作物である文書（ここでは、ＴＸ１～ＴＸ３）の改変に相当するものと考えられるため、著作権保護の観点から、これらの処理の成果物である文書ＸＸ１～ＸＸ３は、ユーザＵ１以外のものから閲覧することができないような形式で保存しておくことが望ましい。

【００４２】

前記テキスト加工処理部５は、前記入出力部１を有する通信端末に搭載するようにしてもよいが、インターネット上にサーバとして配置してもよい。

【００４３】

作業用データベース６は、当該テキスト加工処理部５内の各構成要素５Ａ～５Ｃが処理を進めるために、前記文書ＴＸ１～ＴＸ３などの各データを、整理して蓄積しておくためのデータベースである。最終的に前記文書ＸＸ１～ＸＸ３が得られ、テキスト集合蓄積部４に蓄積されたあと、当該作業用データベース６の蓄積内容は、廃棄することが可能である。

【００４４】

なお、ユーザＵ１がＸＭＬ文書ＸＸ１～ＸＸ３を正常に閲覧するためには、入出力部１のブラウザはＸＭＬ対応のブラウザであることを要する。入出力部１が

搭載しているブラウザが通常のＷｅｂブラウザなど、ＸＭＬ非対応のブラウザである場合には、プラグインソフトを利用して、ＸＭＬ対応の機能を持たせるようにしてもよい。

【００４５】

プラグインソフトは、予め静的に入出力部１に搭載しておくほか、インターネット経由で動的に入出力部１に搭載させることも可能である。

【００４６】

以下、上記のような構成を有する本実施形態の動作について、図２のフローチャートを参照しながら説明する。

【００４７】

図２のフローチャートは、Ｓ１～Ｓ６の各ステップから構成されている。

【００４８】

（Ａ－２）実施形態の動作

ユーザＵ１が入出力部１のブラウザで検索エンジン３にアクセスして所望の検索条件を供給すると（Ｓ１）、検索エンジン３は当該検索条件に適合する文書を、前記テキストデータベース２から検索する（Ｓ２）。

【００４９】

ステップＳ１を実行する際、入出力部１のブラウザに表示される画面は、例えば、図３に示すものであってよい。

【００５０】

図３において、当該画面を構成するウインドウＷＤ１はユーザＵ１からの入力を受け入れるための入力用の領域ＥＲ１と、基本的にユーザＵ１への出力を返すために使用される出力用の領域ＥＲ３に分けられ、入力用の領域ＥＲ１には、フィールドＦＤ１と、ボタンＢＴ１が配置され、出力用の領域ＥＲ２には、フィールドＦＤ２と、画面切り替えボタンＢＴ２～ＢＴ５が配置されている。

【００５１】

このうちフィールドＦＤ１は、ユーザＵ１からの検索キーの入力を受け入れる検索キー入力部である。ここでは、検索キーとして、日付を含む複数のキーワードの入力を許容するものとするが、必要ならば、文書が作成された日付の範囲（

例えば、２００２年６月１１日以降に作成された文書のなかから検索）など、各種の検索条件を柔軟かつ詳細に指定できるようにしてもよい。

【００５２】

検索キー入力部ＦＤ１に入力した検索キーの内容が確定し、ユーザＵ１が「検索開始」ボタンＢＴ１を操作すると、検索エンジン３に、当該検索キーが供給されて検索が実行される。図３の例では、検索キー入力部ＦＤ１に、「Ｚ選手」（野球選手の名前）と、「１５日」と、「ＣＣチーム戦」の３つのキーワードを入力している。

【００５３】

ここでは、当該３つのキーワードに対応する検索結果として、上述した３つの文書ＴＸ１～ＴＸ３が得られたものとする。

【００５４】

ただし本実施形態の場合、単に検索結果である前記文書ＴＸ１～ＴＸ３をフィールド（検索結果出力部）ＦＤ２に表示するのではなく、前記テキスト加工処理部５による処理の結果を、フィールドＦＤ２に表示するため、フィールドＦＤ２に表示が行われるのは、以降の各ステップＳ３～Ｓ６が実行されたあとである。

【００５５】

検索エンジン３による検索の結果として得られた前記３つの文書ＸＴ１～ＸＴ３は、ステップＳ３で、作業用データベース６内のテキスト情報格納テーブルＴＢ１に蓄積される。

【００５６】

テキスト情報格納テーブルＴＢ１の格納内容は、例えば、図４に示すものであってよい。

【００５７】

図４において、当該テキスト情報格納テーブルＴＢ１は、２つの列名（属性）、すなわち、「出典情報」と、「テキスト内容」から構成されている。

【００５８】

検索エンジン３の検索によって得られた文書ＴＸ１～ＴＸ３の数が３であることから、当該テキスト情報格納テーブルＴＢ１の行（タプル）の数も３となって

いる。

【００５９】

図示の例では、出典情報として、各文書ＴＸ１～ＴＸ３の出典である新聞の名称と日付が記述されている。これは人間にとって可読でネットワーク以外の一般社会で使用されるオフラインの出典情報の一例である。必要ならば、このようなオフラインの出典情報に替えて、あるいはオフラインの出典情報とともに、オンラインの出典情報も記述するようにしてもよい。オンラインの出典情報としては、各文書ＴＸ１～ＴＸ３の前記テキストデータベース２上における存在場所を一意に指定する情報、例えば、ＵＲＬ、ＦＱＤＮ、ＩＰアドレスなどを利用することができる。

【００６０】

図４中のテキスト内容から明らかなように、これらの文書ＴＸ１～ＴＸ３は、米国のＰ野球リーグで、野球選手Ｚ（外野手）の属するＢＢチームが、ＣＣチームと対戦した試合における当該Ｚ選手の活躍ぶりを報じる同日付けの新聞記事である。したがって、文書ＴＸ１～ＴＸ３の内容であるテキスト内容は大部分が同じであるが、例えば、Ｂ新聞の記事である文書ＴＸ２ではこの試合でＺ選手が打率を３割４分９厘に落としたことが記載されているのに、Ａ新聞の記事である文書ＴＸ１や、Ｃ新聞の記事である文書ＴＸ３ではそのことに関する記載は存在しないなど、相違点も存在する。

【００６１】

次に、前記主題情報生成部５Ａが、当該テキスト情報格納テーブルＴＢ１の格納内容をもとに、前記主題情報ＴＨ１を生成する（Ｓ４）。

【００６２】

ここでは、上述した要約生成法と代表選択法のうち、要約生成法を用いて主題情報ＴＨ１を生成したものとする。

【００６３】

要約生成法で生成された要約ＴＸＡは、少なくとも、テキスト加工処理部５における処理が終了するまでの間、作業用データベース６内に前記テキスト情報格納テーブルＴＢ１とともに格納しておく。もちろん、必要ならば、テキスト情報

格納テーブルＴＢ１のなかに、要約ＴＸＡの内容を格納しておくための列名を用意してもよい。

【００６４】

このあと、当該要約ＴＸＡに対する各文書ＴＸ１～ＴＸ３の差分情報を抽出する（Ｓ５）。ここでは、前記単位として、節を使用しているため、前記ＸＭＬ形式への変換や、変換後のＸＭＬ文書ＸＸ１～ＸＸ３のテキスト集合蓄積部４への格納などを行ったあと、ユーザＵ１からの出力要求に応じて入出力部１上で前記検索結果出力部ＦＤ２に表示される内容は、例えば、図７に示すようになる。

【００６５】

図７において検索結果出力部ＦＤ２内には、最上部に、前記主題情報ＴＨ１が表示され、その下に、１行おいて、オフラインの出典情報ＯＦ１である「Ａ新聞５月１６日」とＡ新聞５月１６日の記事の当該主題情報ＴＨ１に対する差分情報ＥＨ１、オフラインの出典情報ＯＦ２である「Ｂ新聞５月１６日」とＢ新聞５月１６日の記事の当該主題情報ＴＨ１に対する差分情報ＥＨ２、オフラインの出典情報ＯＦ３である「Ｃ新聞５月１６日」とＣ新聞５月１６日の記事の当該主題情報ＴＨ１に対する差分情報ＥＨ３がそれぞれ表示されている。

【００６６】

文書（例えば、ＸＸ１）のなかから、差分情報（ここでは、ＥＨ１）だけを抽出して、例えば図７に示すように画面表示を行う処理は、前記タグの属性に各単位が差分に該当する旨を指定しておくことによって、入出力部１上のＸＭＬ対応ブラウザ（あるいは、前記プラグインを装備したブラウザ）の機能だけで実行可能である。

【００６７】

文書ＸＸ１～ＸＸ３中の差分に該当する単位は、図５にアンダーラインを付して示した部分である。

【００６８】

図７の画面が入出力部１に表示されると、ユーザＵ１は個々の文書ＸＸ１～ＸＸ３の内容を読まなくても、主題情報ＴＨ１を読むだけで、テキスト集合の主題を正確に認識することができる。主題情報ＴＨ１の文字数は文書ＸＸ１～ＸＸ３

のうちの任意の１文書の文字数とほぼ同程度であるので、個々の文書ＸＸ１～ＸＸ３を読む場合に比べ、ユーザＵ１が読むべき文字数はほぼ１／３程度となる上、各文書ＸＸ１～ＸＸ３の記事内容の異同をユーザＵ１の頭脳などを用いて分析する必要もなく、入出力部１上へ個々の文書ＸＸ１～ＸＸ３のファイルをダウンロードしたり、開いたりするための操作を逐一おこなう必要もない。

【００６９】

このためユーザＵ１は極めて簡単に主題情報ＴＨ１を認識することができる。また、これらの効果は、一般的に、１つのテキスト集合中の文書の数が多くなればなるほど、顕著になる。

【００７０】

図７の画面例は、ユーザＵ１が「主題＆差分情報表示」ボタンＢＴ４を操作して出力要求を行った場合に対応する表示画面であるが、ユーザＵ１が「主題＆参照情報表示」ボタンＢＴ３を操作して出力要求を行ったときには、図６に示す表示画面が表示される。この参照情報は、前記出典情報に等しい。

【００７１】

図６では、前記差分情報ＥＨ１～ＥＨ３が消失して、オフライン出典情報ＯＦ１～ＯＦ３だけが、主題情報ＴＨ１の下に表示されている。

【００７２】

一方、図８は、図７の表示画面上でユーザＵ１がオフライン出典情報ＯＦ３をポインティングデバイスなどを用いて選択したときの表示例を示している。

【００７３】

このとき、主題情報ＴＨ１上では、随所にアンダーラインが表示され、主題情報ＴＨ１の内容のうち当該オフライン出典情報ＯＦ３に対応する前記文書ＴＸ３から得られた情報がどれであるかを直観的に示すことができる。同様に、ユーザＵ１がオフライン出典情報ＯＦ２を選択すれば、アンダーラインが表示されて主題情報ＴＨ１の内容のうち当該オフライン出典情報ＯＦ２に対応する前記文書ＴＸ２から得られた情報を示し、ユーザＵ１がオフライン出典情報ＯＦ１を選択すれば、アンダーラインが表示されて主題情報ＴＨ１の内容のうち当該オフライン出典情報ＯＦ１に対応する前記文書ＴＸ１から得られた情報を示すことができる

。

【００７４】

必要に応じて、図６の画面上でも、オフライン出典情報を選択することによって同様なアンダーラインを表示するようにしてもよい。

【００７５】

このアンダーラインは、前記スタイルシート言語を変更することによって、反転表示や網かけ表示などへ適宜、変更可能である。また、図６～図８における検索結果出力部ＦＤ２上のレイアウトなども、スタイルシート言語の変更に応じて変化する。

【００７６】

図６～図８のいずれの画面を目視した場合でも、ユーザＵ１は、主題情報ＴＨ１を読むことによって、文書ＴＸ１～ＴＸ３（あるいは、ＸＸ１～ＸＸ３）で構成されるテキスト集合の主題を、簡単、かつ確実に認識することが可能である。

【００７７】

必要に応じて、各オフライン出典情報ＯＦ１～ＯＦ３と各文書ＸＸ１～ＸＸ３（あるいは、テキストデータベース２上の各文書ＴＸ１～ＴＸ３）を関連づけておくことにより、オフライン出典情報を選択したときに、当該文書の全文を表示させること等も実行可能である。

【００７８】

（Ａ－３）実施形態の効果

本実施形態によれば、ユーザ（Ｕ１）は、テキスト集合に含まれる個々の文書（例えば、ＴＸ１～ＴＸ３）を読まなくても、当該テキスト集合の主題（例えば、ＴＨ１）を認識することができ、利便性に優れている。

【００７９】

また本実施形態では、個々の文書と主題との差分情報（例えば、ＥＨ１～ＥＨ３）を表示させたり、主題情報のなかのどの部分（単位）が、各文書に対応しているかを表示させることもできるため、ユーザが各文書を対比したり、分析したりする作業を支援することが可能である。

【００８０】

（Ｂ）他の実施形態

上記実施形態にかかわらず、入出力部１の通信端末として、ポインティングデバイス等を備えた一般的なパソコンの替わりにタッチパネル装置を使用したり、専用の通信端末を使用したりすることができる。

【００８１】

また、前記文書ＴＸ１～ＴＸ３およびＸＸ１～ＸＸ３には、単なるテキストデータだけでなく画像データなどが含まれていてもかまわないことはすでに述べた通りである。

【００８２】

なお、上記実施形態では、テキスト加工処理部５は、最終的に文書をＸＭＬ形式（あるいは、テキスト形式）に変換してテキスト集合蓄積部４に蓄積したが、必要に応じて、ＸＭＬ形式（テキスト形式）以外のデータ形式に変換するようにしてもよいことは当然である。

【００８３】

さらに上記実施形態では、ＸＭＬのタグや、属性によって、前記単位が差分に該当する旨を示し、なおかつ再利用可能な形で保存するようにしたが、ＸＭＬのタグや属性以外の方法を用いてこれらの機能を実現してもかまわない。

【００８４】

また、上記実施形態では主題情報ＴＨ１の生成にあたり、上述した要約生成法または代表選択法を用いるものとしたが、これら以外の方法で主題情報を生成するようにしてもかまわない。

【００８５】

例えば、テキスト加工処理部５が自動的に所定の定型的な手順（例えば、検索された複数の文書（例えば、ＴＸ１～ＴＸ３）のなかから単に文字数の最も少ない文書を主題情報とする）で主題情報を決定するようにしてもよい。

【００８６】

もともと、検索エンジン３で検索した時点で文書ＴＸ１～ＴＸ３の間の類似度が十分に高い場合などには、このような単純な方法で選択した文書によっても、テキスト集合の主題を、良好に表現することも可能である。

【００８７】

さらに上記実施形態では、主題情報の生成過程にはユーザＵ１が関与することができず、テキスト加工処理部５側が自動的に生成したが、ユーザＵ１の意思に応じて主題情報を生成することも可能である。

【００８８】

例えば、前記テキスト集合中の任意の１文書をユーザＵ１が主題情報として選択できるようにしてもよい。

【００８９】

この場合、ユーザＵ１の選択に応じて、テキスト加工処理部５が動作し、ユーザＵ１が選択した１文書と他の文書との差分情報などが自動的に得られる。このような構成は、相互に類似した複数の文書間で、共通点や相違点を詳細に整理する必要がある場合に有用である。

【００９０】

また、上記実施形態にかかわらず検索エンジン３は省略可能である。

【００９１】

現実の文書処理の局面では、検索エンジン３で検索しなくても、予め複数の文書（例えば、ＴＸ１～ＴＸ３）が与えられているケースも多いからである。また、文書（例えば、ＴＸ１～ＴＸ３）は必ずしもネットワーク経由で供給されるものである必要はない。例えば、フロッピディスクやＣＤ－ＲＯＭなどの記録媒体に格納された形で供給されたり、あるいは、紙媒体の形で供給されたものがＯＣＲ処理などを経てシステム内に取り込まれる場合もあってよい。

【００９２】

また、上記実施形態では、同じ試合における野球選手Ｚの活躍を報じる同日付けの新聞記事であったため、文書ＴＸ１～ＴＸ３の内容が類似していることが明確に予測できる場合であったが、類似しているか否かが不明な複数の文書に対して本発明を適用してもよい。

【００９３】

その場合、本発明を利用して、文書間の類似度を判定する作業を容易化することが可能になる。

【００９４】

なお、上記実施形態で使用したテキスト情報格納テーブルＴＢ１のスキーマは、上述したものに限定する必要はない。テキスト情報格納テーブルＴＢ１中の列名を他の列名に置換してもよく、テキスト情報格納テーブルＴＢ１中に存在しない列名を追加してもよい。このようなテキスト情報格納テーブルを、必要に応じて、正規化してもよいことは当然である。

【００９５】

さらに、前記作業用データベース６とテキスト集合蓄積部４は、ハードウエア的には必ずしも別個に設ける必要はなく、統合可能である。

【００９６】

また、上記実施形態にかかわらず、前記入出力部１は省略可能である。

【００９７】

例えば、予め与えられたプログラム等にしたがって、検索エンジン３による検索や、テキスト加工処理部５による処理を行い、最終結果である文書（例えば、ＸＸ１～ＸＸ３）を、記録媒体に書き込むこと等で処理が完結するシステムもあり得るからである。

【００９８】

また、上記実施形態では、図３、図６～図８に具体的な表示画面例を示したが、本発明の表示画面の構成は図示したものに限らないことは当然である。

【００９９】

さらに、前記文書ＴＸ１～ＴＸ３は、新聞記事であったが、本発明が対象とする文書が新聞記事にかぎらないことは当然である。

【０１００】

以上の説明では主としてソフトウエア的に本発明を実現したが、本発明はハードウエア的に実現することも可能である。

【０１０１】

【発明の効果】

以上に説明したように、本発明の文書処理装置および方法は、従来よりも、利便性に優れている。

【図面の簡単な説明】

【図１】

実施形態に係る閲覧システムの全体構成例を示す概略図である。

【図２】

実施形態の動作を示すフローチャートである。

【図３】

実施形態の動作を示す表示画面例である。

【図４】

実施形態で使用するテキスト情報格納テーブルの内容例を示す概略図である。

【図５】

実施形態で使用するテキスト情報格納テーブルの内容例を示す概略図である。

【図６】

実施形態の動作を示す表示画面例である。

【図７】

実施形態の動作を示す表示画面例である。

【図８】

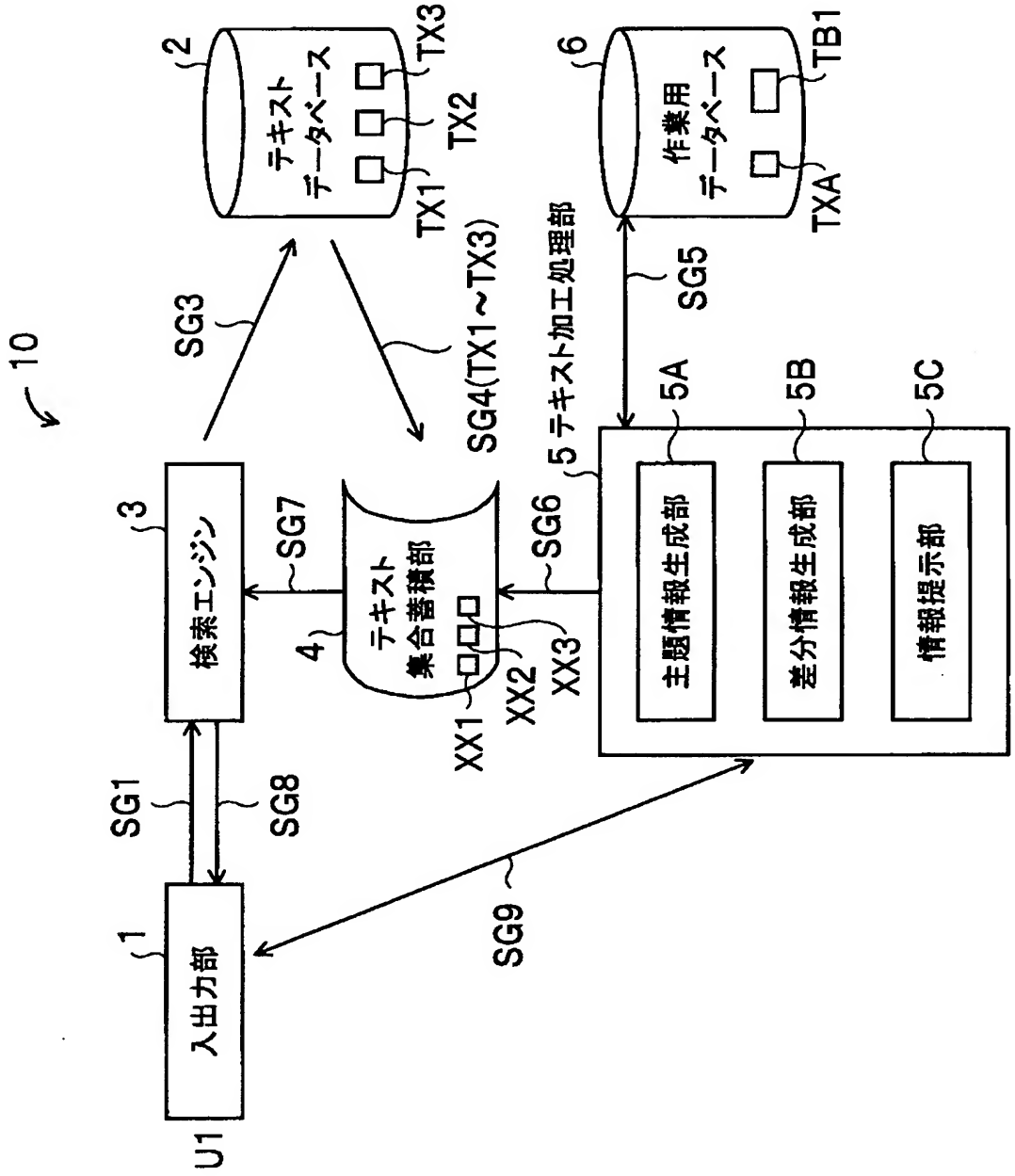
実施形態の動作を示す表示画面例である。

【符号の説明】

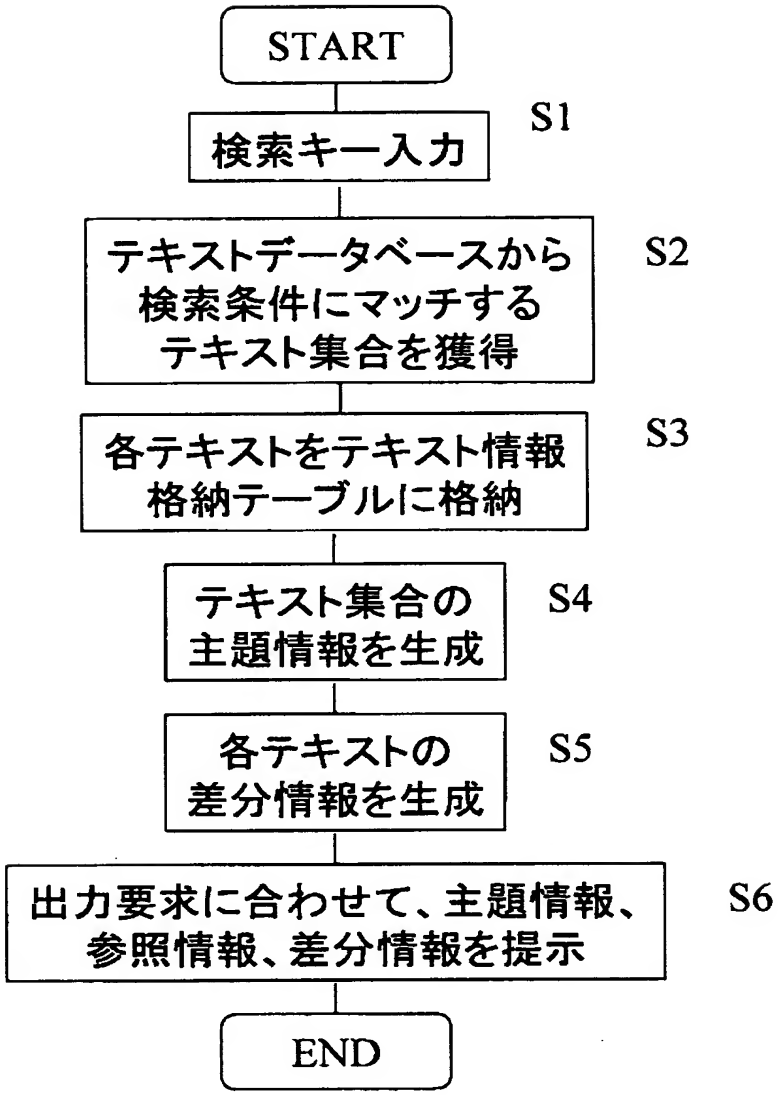
１…入出力部、２…テキストデータベース、３…検索エンジン、４…テキスト集合蓄積部、５…テキスト加工処理部、５Ａ…主題情報生成部、５Ｂ…差分情報生成部、５Ｃ…情報提示部、６…作業用データベース、ＴＢ１…テキスト情報格納テーブル、ＴＸ１～ＴＸ３…文書、ＸＸ１～ＸＸ３…ＸＭＬ文書、ＴＸＡ…要約、ＴＨ１…主題情報（主題）、ＯＦ１～ＯＦ３…オフライン出典情報、ＥＨ１～ＥＨ３…差分情報。

【書類名】　図面

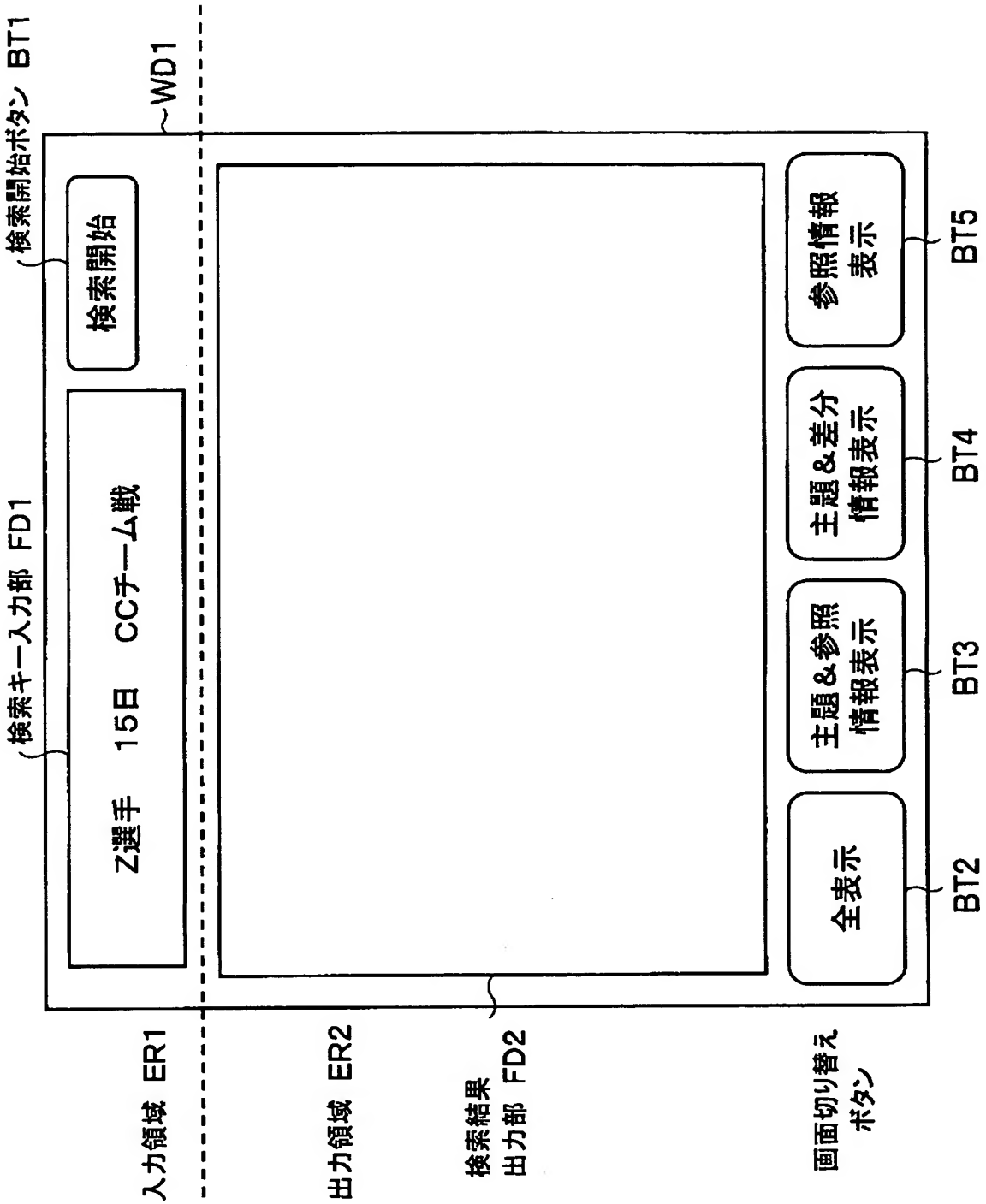
【図１】


10
2
テキスト
データベース
TX3
TX2
TX1
6
作業用
データベース
TB1
TXA
SG3
SG4(TX1～TX3)
SG5
5 テキスト加工処理部
5A
5B
5C
3
検索エンジン
SG7
SG6
4
テキスト
集合蓄積部
XX1
XX2
XX3
主題情報生成部
差分情報生成部
情報提示部
SG1
SG8
SG9
1
入出力部
U1

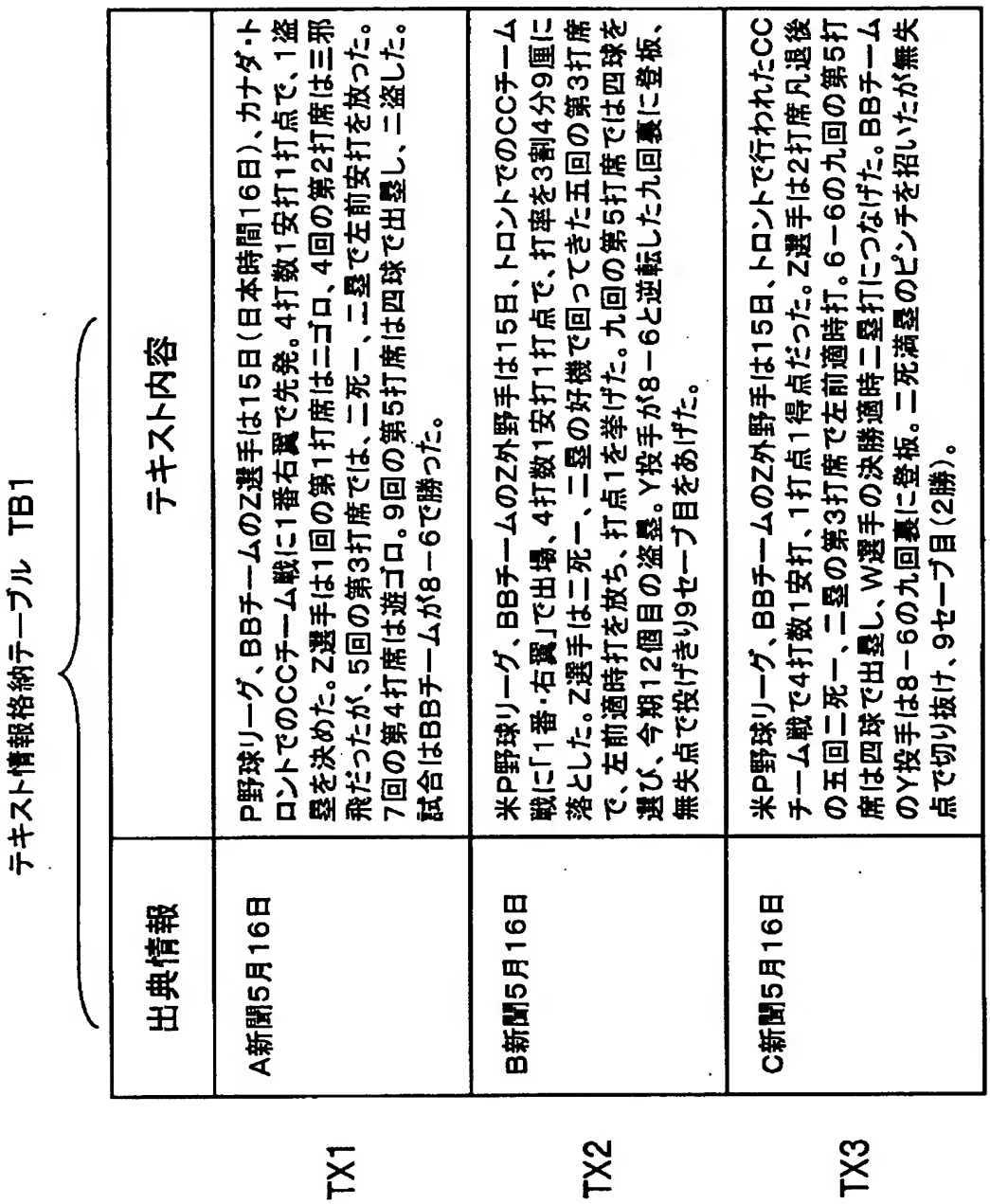
【図２】


START
検索キー入力
S1
テキストデータベースから
検索条件にマッチする
テキスト集合を獲得
S2
各テキストをテキスト情報
格納テーブルに格納
S3
テキスト集合の
主題情報を生成
S4
各テキストの
差分情報を生成
S5
出力要求に合わせて、主題情報、
参照情報、差分情報を提示
S6
END

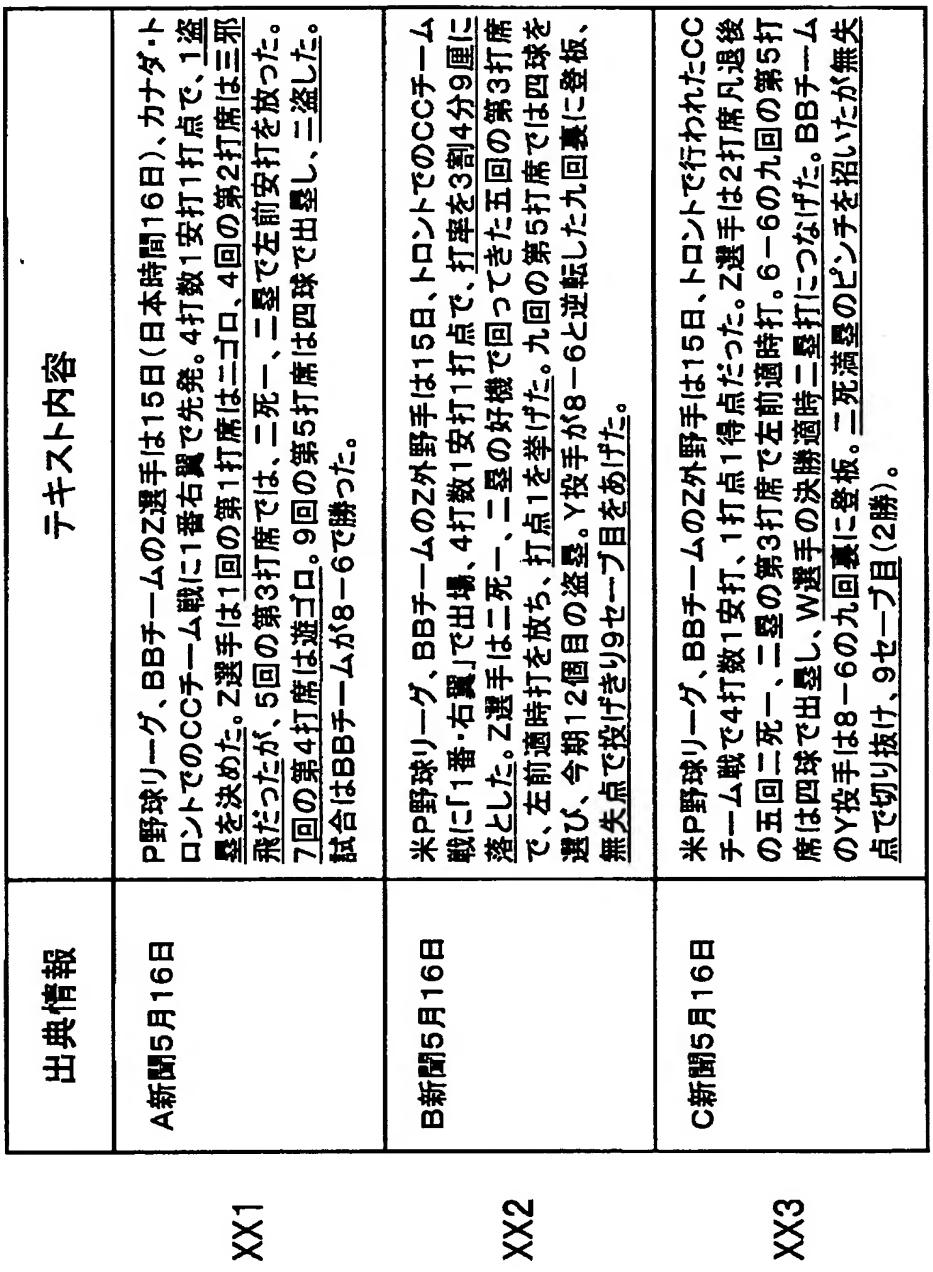
【図３】


検索開始ボタン BT1
WD1
検索キー入力部 FD1
検索開始
Z選手　15日　CCチーム戦
全表示
主題＆参照情報表示
主題＆差分情報表示
参照情報表示
BT2
BT3
BT4
BT5
入力領域 ER1
出力領域 ER2
検索結果出力部 FD2
画面切り替えボタン

【図４】

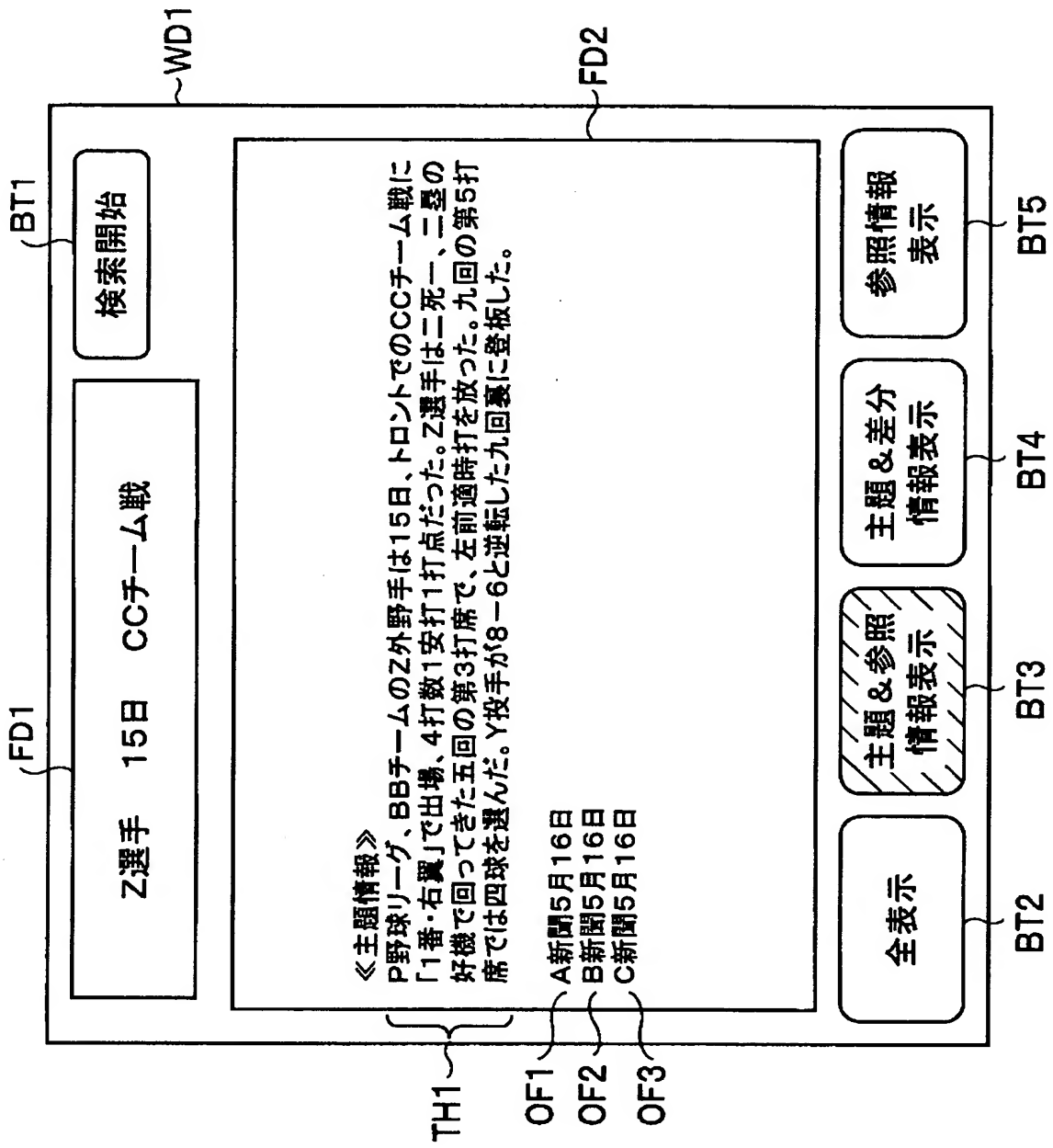

テキスト情報格納テーブル TB1
出典情報
テキスト内容
TX1
A新聞5月16日
P野球リーグ、BBチームのZ選手は15日(日本時間16日)、カナダ・トロントでのCCチーム戦に1番右翼で先発。4打数1安打1打点で、1盗塁を決めた。Z選手は1回の第1打席はニゴロ、4回の第2打席は三邪飛だったが、5回の第3打席では、二死一、二塁で左前安打を放った。7回の第4打席は遊ゴロ。9回の第5打席は四球で出塁し、二盗した。試合はBBチームが8－6で勝った。
TX2
B新聞5月16日
米P野球リーグ、BBチームのZ外野手は15日、トロントでのCCチーム戦に「1番・右翼」で出場、4打数1安打1打点で、打率を3割4分9厘に落とした。Z選手は二死一、二塁の好機で回ってきた五回の第3打席で、左前適時打を放ち、打点1を挙げた。九回の第5打席では四球を選び、今期12個目の盗塁。Y投手が8－6と逆転した九回裏に登板、無失点で投げきり9セーブ目をあげた。
TX3
C新聞5月16日
米P野球リーグ、BBチームのZ外野手は15日、トロントで行われたCCチーム戦で4打数1安打、1打点1得点だった。Z選手は2打席凡退後の五回二死一、二塁の第3打席で左前適時打。6－6の九回の第5打席は四球で出塁し、W選手の決勝適時二塁打につなげた。BBチームのY投手は8－6の九回裏に登板。二死満塁のピンチを招いたが無失点で切り抜け、9セーブ目(2勝)。

【図５】

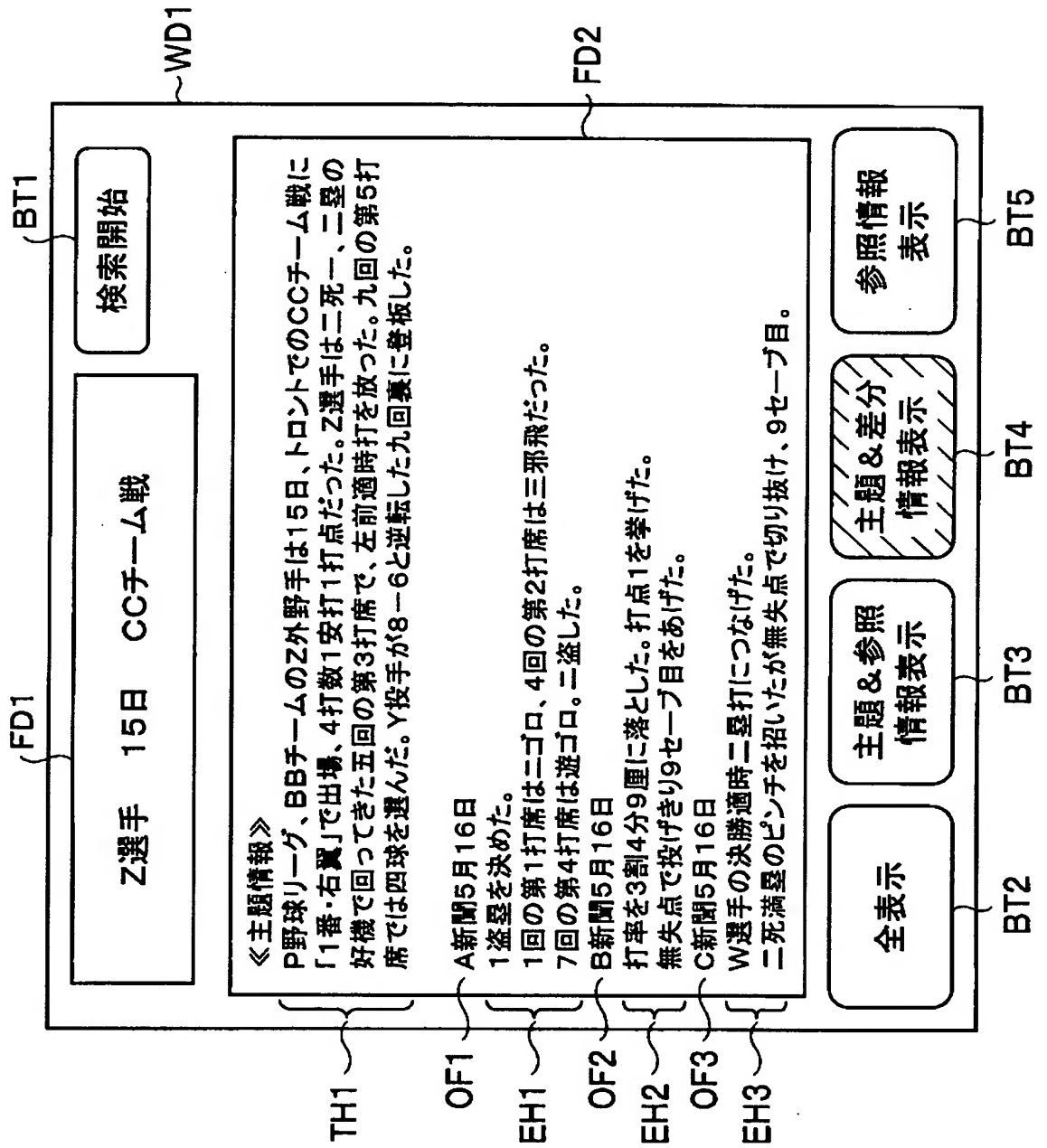


	出典情報	テキスト内容
XX1	A新聞5月16日	P野球リーグ、BBチームのZ選手は15日（日本時間16日）、カナダ・トロントでのCCチーム戦に1番右翼で先発。4打数1安打1打点で、1盗塁を決めた。Z選手は1回の第1打席はニゴロ、4回の第2打席は三邪飛だったが、5回の第3打席では、二死一、二塁で左前安打を放った。7回の第4打席は遊ゴロ。9回の第5打席は四球で出塁し、二盗した。試合はBBチームが8－6で勝った。
XX2	B新聞5月16日	米P野球リーグ、BBチームのZ外野手は15日、トロントでのCCチーム戦に「1番・右翼」で出場、4打数1安打1打点で、打率を3割4分9厘に落とした。Z選手は二死一、二塁の好機で回ってきた五回の第3打席で、左前適時打を放ち、打点1を挙げた。九回の第5打席では四球を選び、今期12個目の盗塁。Y投手が8－6と逆転した九回裏に登板、無失点で投げきり9セーブ目をあげた。
XX3	C新聞5月16日	米P野球リーグ、BBチームのZ外野手は15日、トロントで行われたCCチーム戦で4打数1安打、1打点1得点だった。Z選手は2打席凡退後の五回二死一、二塁の第3打席で左前適時打。6－6の九回の第5打席は四球で出塁し、W選手の決勝適時二塁打につなげた。BBチームのY投手は8－6の九回裏に登板。二死満塁のピンチを招いたが無失点で切り抜け、9セーブ目（2勝）。

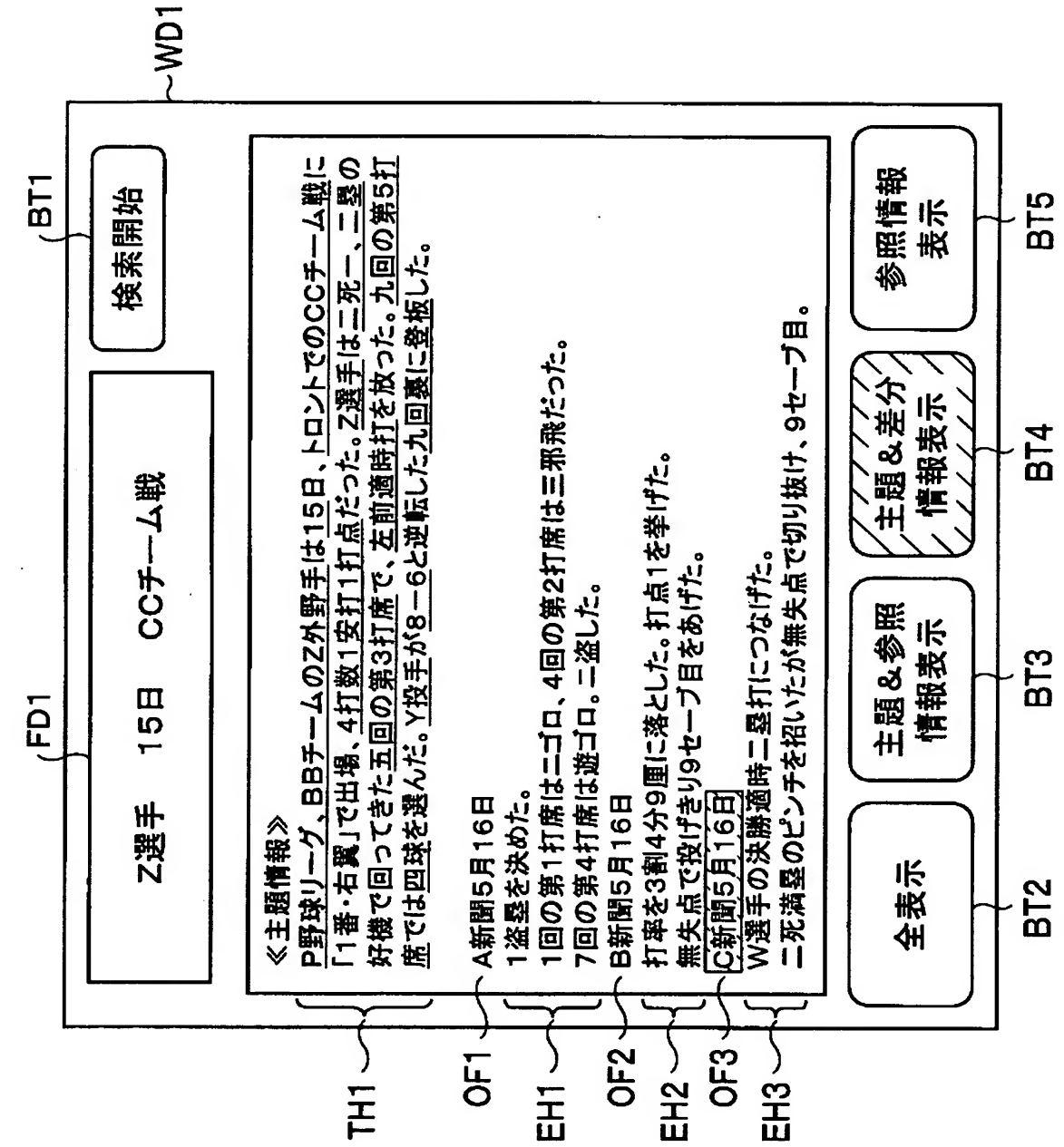
【図６】


WD1
FD2
BT1
検索開始
FD1
Z選手　15日　CCチーム戦
≪主題情報≫
P野球リーグ、BBチームのZ外野手は15日、トロントでのCCチーム戦に「1番・右翼」で出場、4打数1安打1打点だった。Z選手は二死一、二塁の好機で回ってきた五回の第3打席で、左前適時打を放った。九回の第5打席では四球を選んだ。Y投手が8－6と逆転した九回裏に登板した。
A新聞5月16日
B新聞5月16日
C新聞5月16日
TH1
OF1
OF2
OF3
全表示
BT2
主題＆参照情報表示
BT3
主題＆差分情報表示
BT4
参照情報表示
BT5

【図７】


WD1
FD1
Z選手　15日　CCチーム戦
BT1
検索開始
FD2
≪主題情報≫
P野球リーグ、BBチームのZ外野手は15日、トロントでのCCチーム戦に「1番・右翼」で出場、4打数1安打1打点だった。Z選手は二死一、二塁の好機で回ってきた五回の第3打席で、左前適時打を放った。九回の第5打席では四球を選んだ。Y投手が8－6と逆転した九回裏に登板した。
A新聞5月16日
1盗塁を決めた。
1回の第1打席はニゴロ、4回の第2打席は三邪飛だった。
7回の第4打席は遊ゴロ。二盗した。
B新聞5月16日
打率を3割4分9厘に落とした。打点1を挙げた。
無失点で投げきり9セーブ目をあげた。
C新聞5月16日
W選手の決勝適時二塁打につなげた。
二死満塁のピンチを招いたが無失点で切り抜け、9セーブ目。
全表示
主題＆参照情報表示
主題＆差分情報表示
参照情報表示
BT2
BT3
BT4
BT5
TH1
OF1
EH1
OF2
EH2
OF3
EH3

【図８】


WD1
FD1
BT1
検索開始
Z選手　15日　CCチーム戦
≪主題情報≫
P野球リーグ、BBチームのZ外野手は15日、トロントでのCCチーム戦に「1番・右翼」で出場、4打数1安打1打点だった。Z選手は二死一、二塁の好機で回ってきた五回の第3打席で、左前適時打を放った。九回の第5打席では四球を選んだ。Y投手が8－6と逆転した九回裏に登板した。
A新聞5月16日
1盗塁を決めた。
1回の第1打席はニゴロ、4回の第2打席は三邪飛だった。
7回の第4打席は遊ゴロ。二盗した。
B新聞5月16日
打率を3割4分9厘に落とした。打点1を挙げた。
無失点で投げきり9セーブ目をあげた。
C新聞5月16日
W選手の決勝適時二塁打につなげた。
二死満塁のピンチを招いたが無失点で切り抜け、9セーブ目。
全表示
主題＆参照情報表示
主題＆差分情報表示
参照情報表示
BT2
BT3
BT4
BT5
TH1
OF1
EH1
OF2
EH2
OF3
EH3

【書類名】　　要約書

【要約】

【課題】　利便性を高める。

【解決手段】　文字情報を含む複数の文書を要素とする集合に関して処理を行う文書処理装置において、前記集合中の各文書に関し、前記文字情報の共通性を抽出して、前記集合全体に共通の意味内容を表現した文書である共通文書を生成する共通文書生成手段を備える。またこの文書処理装置において、前記共通文書生成手段が、前記集合中の複数の文書をもとに所定の生成手順を実行して、新たな文書として前記共通文書を生成するか、または、予め文字情報に共通性のある文書を選んで前記集合を構成した上で、前記集合中の複数の文書のなかから所定の選択手順に応じて１つの文書を選択し、選択した当該文書を前記共通文書とすることで前記共通文書を生成することは好ましい。

【選択図】　図１

出 願 人 履 歴 情 報

識別番号　　〔０００００２９５〕

１．変更年月日　　１９９０年　８月２２日
［変更理由］　新規登録
住　所　東京都港区虎ノ門１丁目７番１２号
氏　名　沖電気工業株式会社